L2 transfer of stress, tones, and intonation from Mandarin: A case study

Una Y. Chow
University of Calgary

Abstract

This study examined the prosodic patterns of Mandarin, Cantonese, and English in order to address the question: Will a native speaker of Mandarin acquire Cantonese intonation more easily than English intonation? According to the Markedness Differential Hypothesis (Eckman 1997), second language (L2) features that are universally rarer than the first language (L1) features will create difficulty for L2 acquisition. English has word stress, Cantonese has lexical tones, and Mandarin has both. English has more variation in word stress patterns than Mandarin, and Cantonese has more lexical tones than Mandarin. The prediction was that a Mandarin speaker would have difficulty in acquiring English stress and Cantonese tones.

In a field study, I elicited speech samples from a female, adult native speaker of Mandarin who learned Cantonese and English from age 5-6. My pitch analysis of her speech revealed near native-like intonation patterns in English. In Cantonese, however, her declarative questions reflected an overall raise in pitch range, characteristic of her Mandarin questions. My results demonstrated that the consultant showed more difficulty in her acquisition of the native intonation of Cantonese than that of English. The implication is that lexical tones interfere with L2 intonation more so than word stress, because both lexical tones and intonation rely on fundamental frequency (F0) as a primary cue.

1. Introduction*

In speech, intonation is meaningful variation in the pitch of the voice. It functions as a grammatical marker for different types of sentences, such as statements (as in 1a) and questions (as in 1b, c). For example, a yes/no question is a type of question that expects either a 'yes' or 'no' response (1c). A declarative question is a type of yes/no question that has the same syntactic structure as a statement. A declarative echo question (1b) repeats part or all of what the speaker has just heard.

(1) a. John is reading a book. (statement)
b. John is reading a book? (declarative echo question)
c. Is John reading a book? (yes/no question with subject-aux inversion)

Because declarative (echo) questions lack subject-auxiliary inversion and an overt question marker, intonation is used to distinguish between statements and declarative questions in speech. However, intonation systems differ from language to language. For example, English signals echo questions with a rising intonation at the end of the utterance (Wells 2006), whereas Mandarin raises the pitch range of the entire utterance (Peng *et al.* 2005). Because English's question intonation differs from Mandarin's question intonation, native English speakers may fail to recognize Mandarin's question intonation. They may misperceive the raise in pitch as anger, for example, since anger is also expressed with increased pitch and pitch range (Williams & Stevens 1972). Thus, cross-linguistic differences pose a challenge for L2 learners.

According to the Markedness Differential Hypothesis (Eckman 1997), L2 features that are universally rarer than the L1 features will create difficulty in L2 learning. Many L2 studies have compared native Mandarin speakers to bilingual Mandarin-English speakers (Marinova-Todd *et al*. 2010, among others) or to bilingual Mandarin-Cantonese speakers (Chen *et al*. 2004, among others). Few studies, if any, have compared the acquisition of two L2s by the same speaker. To gain a better understanding on native language interference on the acquisition of L2 intonation, this case study examines the prosodic patterns of Mandarin, Cantonese, and English of a trilingual speaker in order to address the question: Will a native speaker of Mandarin acquire Cantonese intonation easier than English intonation?

English has word stress, Cantonese has lexical tones, and Mandarin has both. Between the stress languages, English has more variation in stress patterns (*e.g.*, as shown by the capitalized syllables in the following pairs: *phoNEtic* [fə.'nɛ.tɪk] vs. *PHOnetician* [ˌfɑ.nə.'tʰɪ.ʃən], or *CAnada* ['kʰæ.nə.də] vs. *baNAna* [bə.'næ.nə]) than Mandarin (*e.g.*, 妹妹 *MEI4 mei* [mei^{51}.mei] 'younger sister'). Between the two tone languages, Cantonese has six contrastive tones, while Mandarin has four. The prediction is that a Mandarin speaker would

* I would like to extend my sincerest gratitude to Dr. Stephen Winters and Dr. Darin Flynn for their encouragement and mentorship on this project and for their helpful comments on the earlier drafts of this paper. Also, I would like to thank Choon Foong Chow-Wah and Angela Choy for their assistance in this project. This project would not be possible without the commitment and tremendous patience of my wonderful language consultant.

have difficulty in acquiring English stress and Cantonese tones. Given that word stress is perceived based on F0, intensity, and duration in English, and that lexical tone is perceived primarily based on F0 in Cantonese, both of which interact with intonation (Ladd 2008), the prediction is that there would be L1 interference and that there would be some degree of intonation transfer from the speaker's L1 to L2.

2. Background

2.1. Mandarin and Cantonese lexical tones

Mandarin has a tonal system of four contrastive tones and one neutral tone, as shown in Table 1. The contrastive, or specified, tones are high-level (mT1), rising (mT2), low-fall-rise (mT3), and falling (mT4). Using Chao's (1930) five-scale tonal system, with 1 the lowest and 5 the highest point in a speaker's F0 range, mT1 to mT4 are represented as [55], [35], [214], and [51], respectively. The third tone, mT3, has two allotones: [21] and [214]. The neutral tone, written without a tone number (*e.g.*, *ma* '(*question particle*)'), is unspecified and assimilates the tone from the immediately preceding syllable that carries a specified tone.

In Mandarin, every syllable has a lexical tone, including the four specified tones and one unspecified tone. In addition, the first syllable of a word must have a specified or full tone (mT1, mT2, mT3, or mT4). The unspecified, neutral tone occurs only on non-word-initial syllables.

Tone	**mT1**	**mT2**	**mT3**	**mT4**
Number	1	2	3	4
Shape	High-level	Rising	Low-fall(-rise)	Falling
Pitch level	55	35	21(4)	51
Phonological feature	H	LH	L	HL
Example	*ma1* 'mother'	*ma2* 'hemp'	*ma3* 'horse'	*ma4* 'to scold'

Table 1. The four contrastive tones of Mandarin.
Phonological feature: H/L = [+/- upper register] (Yip 1980, Pulleyblank 1986).

Cantonese has a tonal system of six contrastive tones, as shown in Table 2. Unlike Mandarin, every syllable has a specified tone. The contrastive, or specified, tones are high-level (cT1), high-rise (cT2), mid-level (cT3), low-fall (cT4), low-rise (cT5), and low-level (cT6). Using Chao's five-scale tonal system, cT1 to cT6 are represented as [55], [25], [33], [21], [23], and [22], respectively. According to Bauer & Benedict (1997:131), cT1 has an allotone that is high-falling [53], but most Cantonese speakers in Hong Kong produce cT1 with the [55] tone.

Tone	cT1	cT2	cT3	cT4	cT5	cT6
Number	1	2	3	4	5	6
Shape	High-level	High-rise	mid-level	Low-fall	Low-rise	Low-level
Pitch level	55	25	33	21	23	22
Phonological feature	Hh	Hlh	Hl	Ll	Llh	Lh
Example	*fan1* 'point'	*fan2* 'flour'	*fan3* 'training'	*fan4* 'grave'	*fan5* 'raise'	*fan6* 'portion'

Table 2. The six contrastive (non-checked) tones of Cantonese. Shape and pitch level (source: Flynn 2003). Phonological feature: H/L = [+/-upper register], h/l = [+/- raised pitch] (Yip 1980, Pulleyblank 1986).

There are two types of lexical tones in Cantonese: non-checked tones which occur on open syllables and checked tones which occur on closed syllables with an unreleased voiceless stop in the coda (*i.e.*, [pˀ], [tˀ], or [kˀ]). The six unchecked tones are cT1 to cT6. There are three checked tones (cT7, cT8, and cT9), which have the same tonal shape as cT1, cT3, and cT6, respectively, but occur on closed syllables. For example, 食 *sik6* [sɪkˀ^{22}] 'eat' has a checked [22] tone (cT9), whereas 事 *si6* [siː^{22}] 'matter' has a non-checked [22] tone (cT6). Checked tones are usually perceived as shorter in duration because the vowel which carries the tone (and sonority) is cut off by the obstruent coda. In contrast, although the majority of syllables in Cantonese has a consonant-vowel (CV) structure, the vowel in these syllables is lengthened in favour of a binary foot, *e.g.*, *si6* [siː^{22}]. Therefore, non-checked tones are perceived as longer.

Because there are six contrastive tones in Cantonese and each one can occur on an independent syllable, Cantonese has more monosyllabic words than Mandarin. In addition, because a high frequency of words in Cantonese is bimoraic (*i.e.*, CVː or CVC), the rhythmic pattern of spoken Cantonese is perceived as syllable-timed, with approximately the same amount of time given to each syllable.

2.2. Mandarin and English word stress

Typically, native Mandarin speakers perceive stress contrasts in trochees with a full tone on the first syllable and a neutral tone on the second syllable (Peng *et al.* 2005, Duanmu 2007). This stress pattern appears in words with either a reduplicated syllable (*e.g.*, 妹妹 *mei4 mei* ['mei^{51}.mei] 'sister') or a suffix (e.g., 他們 *ta1 men* ['tʰa^{55}.mən] 'they'). Lexical minimal pairs based on stress patterns alone are rare in Mandarin.

English has a more complicated word stress system than Mandarin. Normally, the stress is on the penult, but it can fall on a different syllable depending on the prosodic context. Lexical minimal pairs based on stress patterns exist in English, *e.g.*, the noun *permit* ['pʰɻ̩.mɪt] vs. the verb *permit* [pɻ̩.'mɪt]. The rhythmic pattern of spoken English is perceived as stress-timed, with a fairly equal amount of time spaced between consecutive stressed syllables. The rhythmic pattern of spoken Mandarin, on the other hand, is perceived as syllable-timed (Mok 2009).

2.3. Intonation of declarative questions in Mandarin, Cantonese, and English

Mandarin and Cantonese have at least two types of yes/no questions: A-not-A questions and declarative questions. A-not-A questions are translated into *do*-questions in English, as the examples in (2) show.

(2) a. *ni3 hui4-bu4-hui4 juo4 shi1?* (Mandarin)
you know-not-know create poem
'Do you know how to write a poem?'

b. *nei5 sik1-m4-sik1 jok3 si1?* (Cantonese)
you know-not-know create poem
'Do you know how to write a poem?'

Because A-not-A questions are syntactically marked by the A-not-A construction, variability in intonation does not affect the interrogative mood of the sentence. Therefore, there is no motivation to have specified intonation patterns to signal this type of questions.

Declarative questions are not syntactically marked as questions. They appear in the same word order as a statement, as shown in (3).

(3) a. He has a hundred percent. (statement)
b. He has a hundred percent? (declarative echo question)

Thus, there is motivation for distinct intonation patterns between statements and declarative questions. In Mandarin, declarative echo questions are typically produced with a raised pitch range (Yuan *et al.* 2002). That is, the pitch range of the speaker on the entire utterance is relatively higher than that on the statement. In English (Wells 2006) and Cantonese (Gu *et al.* 2005), however, declarative echo questions are realized with a rise in pitch or a high boundary tone (H%) at the end of the utterance only.

3. Methods

3.1. The language consultant

The language consultant for this case study was a female native speaker of Mandarin, born and raised in Singapore. She moved to Canada at age 22 and lived there for more than ten years. Mandarin was the first language that she learned and also the primary language spoken at home when she was growing up. At age five, she began to learn Cantonese, mostly from watching television. From age six, she learned English formally in school, along with Mandarin. At the time of the study, she was speaking English at home with her family, but continued to speak Mandarin and Cantonese with her friends, occasionally code-switching among Mandarin, Cantonese, and English. She could speak and read all three languages fluently. She could read traditional Chinese characters. The dialects of her Mandarin, Cantonese, and English resembled Huayu, Hong Kong Cantonese, and British English, respectively.

3.2. Elicitation sessions

The language consultant participated in 14 weekly sessions of two hours each between January and April 2011. All elicitation sessions were conducted by the author in a quiet room in the consultant's house, except for one which was at my home. In each session, the consultant performed two or three reading tasks. For each task, she was first presented with a list of words, phrases, or sentences on a sheet of paper in either traditional Chinese characters for eliciting Mandarin or Cantonese, or in English for eliciting English. She was then asked to replace any words or phrases that did not sound natural to her or that she could not say. After the consultant had familiarized herself with the list, she was asked to say each token aloud as naturally as she could, as if she was speaking to her friends or family.

Each reading task was repeated twice. In the first iteration, I manually transcribed the prosodic features (*i.e.*, tones, stress, and F0 contours) of the utterance by ear. The consultant spoke more naturally when she was not being recorded. In the second iteration, I audio-recorded the readings with Praat (Boersma & Weenink 2010) on a MacBook Pro for Tones and Break Indices (ToBI) annotation (Beckman *et al.* 2005) of the prosodic features after the session. During the recording phase, the consultant spoke each token twice in sequence. A token was re-recorded if the consultant misread it, or if there was noticeable background noise, such as phone ringing. In general, the ToBI annotation was used for the final analysis. However, where Praat failed to show the F0 contours correctly (*e.g.*, due to low tones, vowel devoicing, or perturbation), the manual transcription aided in determining the final transcription, if possible.

Corresponding tokens from two or all three languages were elicited. For example, the same Chinese tokens 媽 'mother', 麻 'hemp', 馬 'horse', and 罵 'to scold' were used to elicit both Mandarin and Cantonese tones; place names with similar pronunciation, such as 密西西比 'Mississippi', were used to elicit Mandarin and English stress; and sentences with equivalent meanings, such as 你會不會作詩? (Mandarin), 你識唔識作詩? (Cantonese), and 'Do you know how to write a poem?' (English), were used to elicit intonation patterns of all three languages. Several reading tasks were repeated at later sessions to verify the accuracy and consistency of earlier readings, including the Mandarin and Cantonese tones, and the statement-and-question pairs in different sentence orders.

3.3. Stimuli

The stimuli were prepared between sessions based on the prosodic patterns of the tokens elicited from the previous session(s). Over a thousand speech samples were elicited, including declarative questions, A-*not*-A yes/no questions and *wh*-questions. However, only part of the data was analyzed and reported in this paper, as follows.

1. Lexical tones
    a. a complete paradigm of the four contrastive Mandarin tones, and
    b. three complete paradigms of the six contrastive Cantonese tones
2. Word stress
    a. English disyllabic words with trochaic or iambic stress patterns, and
    b. Mandarin disyllabic words with a full tone and a neutral tone
3. Intonation
    a. Mandarin, Cantonese, and English cognates or loan words, and
    b. Mandarin, Cantonese, and English statements and declarative (echo) questions.

To investigate the effect of Cantonese tones on sentence intonation, the Cantonese stimuli included a set of statement-and-question pairs, each ending with a different Cantonese tone, as shown in (4). The corresponding Mandarin stimuli were similar expressions in Mandarin, as shown in (5). For corresponding English stimuli, the English translation of the Cantonese sentences was used. To investigate the effect of Mandarin tones on sentence intonation, the stimuli were created using the same method as above but with the four contrastive Mandarin tones.

(4) Declarative questions, ending with the Cantonese syllable 'fan':

a. *keoi5 jau5 jat1 baak3 fan1?* (佢有一百分?)
he have one hundred point
'He has a hundred percent?'

b. *keoi5 soeng2 sik6 loeng4.fan2?* (佢想食涼粉?)
he want eat jello
'He wants to eat jello?'

c. *keoi5 jiu3 heoi3 sau6 fan3?* (佢要去受訓?)
he must go receive training
'He must go for training?'

d. *keoi5 soeng2 heoi3 soeng5 fan4?* (佢想去上墳?)
he want go up grave
'He wants to visit the grave?'

e. *keoi5 dak6.bit6 hing1.fan5?* (佢特別興奮?)
he especially excited
'He is especially excited?'

f. *keoi5 soeng2 jiu3 jat1 fan6?* (佢想要一份?)
he want have one portion
'He wants to have a portion?'

(5) Equivalent set of questions in (4), uttered in Mandarin:

a. *ta1 you3 yi1 bai3 fen1?* (他有一百分?)
he have one hundred point
'He has a hundred percent?'

b. *ta1 xiang3 chi1 liang2.fen3?* (他想吃涼粉?)
he want eat jello
'He wants to eat jello?'

c. *ta1 yao4 qu4 shou4 xun4?* (他要去受訓?)
he must go receive training
'He must go for training?'

d. *ta1 xiang3 qu4 shang4 fen2?* (他想去上墳?)
he want go up grave
'He wants to visit the grave?

e. *ta1 te4.bie2 xing1.fen4?* (他特別興奮?)
he especially excited
'He is especially excited?'

f. *ta1 xiang3 yao4 yi1 fen4?* (他想要一份?)
he want have one portion
'He wants to have a portion?'

In order to rule out effects from sentence order, the statement-and-question pairs were presented in three different sentence orders, as shown in (6), each in a separate session.

(6) a. Statement followed by the question (*S-Q*)
S: He wants to have a share.
Q: He wants to have a share?

b. Question followed by the answer (*Q-A*)
Q: He wants to have a share?
A: Yes, he wants to have a share.

c. Statement, followed by the question and answer (*S-Q-A*)
S: He wants to have a share.
Q: He wants to have a share?
A: Yes, he wants to have a share.

3.4. Transcription system

To assist in identifying any tone and intonation transfer from Mandarin to Cantonese, a labeling system is necessary for annotating lexical and boundary tones of both languages, in particular, focal prominence signaled by raised pitch range over a localized or the entire part of the utterance. Wong *et al.* (2005) proposed a ToBI system, C_ToBI, for annotating Hong Kong Cantonese, and Peng *et al.* (2005) proposed M_ToBI for annotating Pan-Mandarin. Since I needed to identify potential prosodic transfer from Mandarin to Cantonese, I combined elements from C_ToBI and M_ToBI systems into one system, called MC_ToBI (see

Appendix A: Tables A1-A6). C_ToBI formed the basis of MC_ToBI. It adhered to the traditional lexical tone numbers instead of those proposed for C_ToBI. Since I also needed to identify potential prosodic transfer from Mandarin to English, I used break indices that were compatible with the indices from the Mainstream American English ToBI system (MAE_ToBI) (Beckman *et al*. 2005). Then I incorporated the Mandarin tones, stress levels, and tags for backdrop pitch range effects from M_ToBI into MC_ToBI. A *gloss* tier and defined uses of the *misc* tier were also added. C_ToBI doubled the first number of all the contour non-checked tones to distinguish them from the checked tones, *e.g.*, [35] → [335] for non-checked and [35] for checked tone. This numbering system could create confusion since [335] could refer to [33] or [35]. Instead, I marked contour tones on checked syllables with a '-' after the tone number, *e.g.*, *sik1* [53-].

4. Results

4.1. Comparison of Mandarin and Cantonese lexical tones

Table 3 shows a representation of the consultant's F0 range and duration of the four contrastive Mandarin tones, with the consultant's production of *ma*. Similarly, Table 4 shows a representation of the consultant's F0 range and duration of the six contrastive Cantonese tones, with the consultant's production of *si*. The values are averages of both readings.

Tone	**mT1**	**mT2**	**mT3**	**mT4**
Token	媽 'mother'	麻 'hemp'	馬 'horse'	罵 'scold'
F0 (Hz)	192	129 → 231	175 → 106	207 → 121
Duration (ms)	650	830	820	550

Table 3. F0 and duration of the consultant's production of *ma* in Mandarin.
The F0 values of mT2, mT3, and mT4 show a rise or fall in pitch.

Tone	**cT1**	**cT2**	**cT3**	**cT4**	**cT5**	**cT6**
Token	詩 'poem'	使 'to cause'	試 'to try'	時 'time'	市 'city'	事 'matter'
F0 (Hz)	200	165 → 222	178	179 → 141	172 → 120	175 → 125
Duration (ms)	930	925	780	865	920	1065

Table 4. F0 and duration of the consultant's production of *si* in Cantonese.
The F0 values of cT2, cT4, cT5, and cT6 show a rise or fall in pitch.

In isolation, these tones are 2 to 3 times longer than in connected speech. Overall, the consultant spoke the Cantonese tones with longer duration than the Mandarin tones. This is expected because the consultant was more familiar with her native language, so she tended to speak Mandarin faster than Cantonese during elicitation. Her shortest Mandarin tone, mT4, is nearly half the duration of her longest Cantonese tone, cT6. The short duration of her mT4 is consistent with the general production of this tone. Her unusually long cT6 suggests that she might be having difficulty in producing this tone.

As for the F0 values of these tones, they are approximately the same in both languages, ranging from 120 to 240 Hz in isolation, with the exception of mT3 which drops to 106 Hz. In Northern China, mT3 is normally produced as a [214] tone, but the consultant

produced it as a [21]. The [21] allotone matches cT4 in shape, thus, facilitating mapping between these two tones.

The last two Cantonese tones were problematic for the consultant: cT5, a low-rise [23] tone, and cT6, a low-level [22] tone. She produced both incorrectly as a falling tone, with similar F0 value and pitch movement as cT4, a low-fall [21] tone. Further analysis of the additional readings of Cantonese tokens ending with *ji* and *fan*, shown in Table 5, confirmed that the consultant was consistent in producing these cT5 and cT6 errors. When I compared the Cantonese tone letters of these tokens with the tone letters of the Mandarin pronunciation, an interesting pattern emerged. In five out of six cases, cT5 and cT6 both corresponded to mT4 in Mandarin. Interestingly, all of the cT5's and cT6's falling-tone errors were produced on tokens which had an mT4 tone in Mandarin, as shown in Table 5.

Token	**Cantonese** (Jyutping)	**Mandarin** (Pinyin)	**Consultant** F0 (Hz)	**Token**	**Cantonese** (Jyutping)	**Mandarin** (Pinyin)	**Consultant** F0 (Hz)
市 'city'	si5	shi4	172 → 120	事 'matter'	si6	shi4	175 → 125
耳 'ear'	ji5	er3	176	二 'two'	ji6	er4	175 → 126
憤 'angry'	fan5	fen4	182 → 132	份 'portion'	fan6	fen4	152 → 128

Table 5. F0 of the consultant's production of *si*, *ji,* and *fan* in cT5 and cT6. The corresponding tokens in Mandarin are included to show the correlation between cT5/cT6 production errors and mT4.

As the prediction holds, a native Mandarin speaker would have difficulty in acquiring Cantonese tones because the latter has more contrastive tones. The mapping in Table 6 shows that cT1, cT2, and cT4 have corresponding tones in Mandarin, whereas cT3, cT5, and cT6 do not.

Cantonese	**cT1**	**cT2**	**cT3**	**cT4**	**cT5**	**cT6**	GAP
	[55]	[25]	[33]	[21]	[23]	[22]	
	↑	↑		↑			
	[55]	[35]		[21]			[51]
Mandarin	**mT1**	**mT2**	GAP	**mT3**	GAP	GAP	**mT4**

Table 6. Mapping of Mandarin tones to Cantonese tones, based on pitch level.

In one of the three readings of *si*, *ji,* and *fan*, the consultant produced cT3 as a cT6 (*i.e.*, [168 → 134] Hz). However, in general, she produced cT3 correctly in connected speech. The persistent problematic tones were cT5 and cT6.

One possible explanation as to why cT5 and cT6 were produced incorrectly as a falling tone is that the consultant had linked cT5 and cT6 to mT4, a falling tone. Since mT4 is the only Mandarin tone that had not yet been linked to a Cantonese tone, it is potentially available for mapping onto cT5 and cT6. This is shown in Table 7.

Cantonese	**cT1**	**cT2**	**cT3**	**cT4**	**cT5**	**cT6**
	[55]	[25]	[33]	[21]	[23]	[22]
	↕	↕	Acquired	↕	↑	↑
	[55]	[35]		[21]	[51]	[51]
Mandarin	**mT1**	**mT2**		**mT3**	**mT4**	**mT4**

Table 7. The consultant's acquisition of Cantonese tones. One possibility is to map cT5 and cT6 onto mT4.

Another possible explanation relates to the phenomenon that many cognates or loan words with a cT6 in Cantonese have a corresponding mT4 tone in Mandarin, as shown in Table 8. This tonal relationship suggests that native speakers of Mandarin and of Cantonese may perceive cT6 and mT4 as the same tone.

	Chinese Token	Cantonese (Jyutping)	Mandarin (Pinyin)
bus	巴士	baa1 **si6**	ba1 **shi4**
Shanghai	上海	**soeng6** hoi2	**shang4** hai3
Mississippi	密西西比	**mat6** sai1 sai1 bei2	**mi4** xi1 xi1 bi3

Table 8. Tone 6 in Cantonese with corresponding tone 4 in Mandarin, in cognates or loan words.

If cT6 maps onto mT4, cT5 becomes the only tone without an association. Since cT6 is very similar in pitch level to cT5 (*i.e.*, [22] vs. [23] or Lh vs. Llh), it is plausible that cT5 loses its [-raised] (l) pitch and assimilates to cT6, as shown in Table 9.

Cantonese	**cT5**	←	**cT6**
	~~**Llh**~~		**L** h
	L h		↑
			H**L**
Mandarin			**mT4**

Table 9. The consultant's acquisition of Cantonese tones.
Another possibility is mapping mT4 to cT6 and then assimilating cT5 to cT6.

A third possibility is that, phonologically, cT4, cT5, and cT6 all have a low register (*i.e.*, [-upper] feature) with different pitch levels (*i.e.*, [±raised] feature). Perhaps the similarity in register makes these low tones difficult for the consultant to produce. Apparently, she was able to perceive the different [±raised] pitches on a high register but not on a low register. This is similar to the problem that a Cantonese speaker with a cochlear implant encounters when trying to distinguish low tones (Ciocca *et al.* 2002). The F0 values of the consultant's cT4, cT5, and cT6 in the production of *si* shown in Table 4 above (*i.e.*, [179 → 141], [172 → 120], and [175 → 125] Hz, respectively) support the analysis that she was merging all three tones together, as shown in Table 10.

Cantonese	**cT4**		**cT5**		**cT6**	
	[21]		[23]		[22]	
	L	↔	L	↔	L	[-upper]
	↕					
	L					[-upper]
	[21]					
Mandarin	**mT3**					

Table 10. The consultant's acquisition of Cantonese tones. A third possibility is that the [raised] feature is ignored on Cantonese tones with a low register (cT4, cT5, and cT6).

4.2. Comparison of Mandarin and English word stress

A phonological analysis of the speech samples from the English reading tasks shows that, in general, the consultant preferred trochees over iambs. She was presented with disyllabic words that were either a trochee or an iamb (or potentially either), but she tended to place the main stress on the first syllable more frequently than the second, as shown in Table 11. This persistent pattern resembles a common stress pattern in Mandarin which appears on disyllabic words as well: trochees with a full tone followed by a neutral tone, forming a heavy and light binary foot. Samples from the consultant's Mandarin reading tasks reveal the same trochaic pattern, as shown in Table 12. The high-low F0 and intensity patterns in English (Table 11) and Mandarin (Table 12) closely resemble each other, which suggests that the consultant may be using the same acoustic cues to signal stress in both languages. It also suggests that her Mandarin stress pattern is influencing her English.

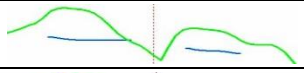
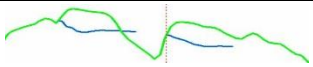
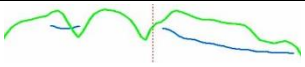
Token	Stress (Target)	Stress (consultant)	F0 (blue), Intensity (green)
southern	['sʌ.ðəɹn]	['saʊf.tən]	
southeast	[saʊθ.'ist]	['saʊθ.ˌist]	
sixteen	[ˌsɪk.'stin] or ['sɪks.ˌtin]	['sɪks.ˌtʰin]	

Table 11. The stress patterns of the consultant's production of English disyllabic words.
The target stress patterns (source: Merriam-Webster Dictionary m-w.com).
The target pronunciation (source: m-w.com, converted to IPA by the author).

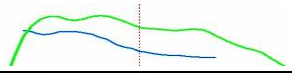
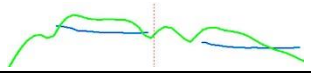
Token	Mandarin (Pinyin)	IPA	Stress	F0 (blue), Intensity (green)
妹妹 'sister'	mei4 mei	[mei^{51}.mei]	heavy - light	
清楚 'clear'	qing1 chu	[tɕʰiŋ^{55}.tʂʰu]	heavy - light	

Table 12. The stress patterns of the consultant's production of Mandarin disyllabic words.

Table 13 shows that the consultant's preference for a trochaic binary foot extends to place names. Although the consultant moved the primary stress from the third syllable to the first syllable in *Mississippi*, she placed a secondary stress on the third syllable to preserve the trochee and binary foot. Since the only prosodic element that was affected is the stressed

foot (Σ), it suggests that the consultant prefers not only a left-headed foot but also a left-headed word. She prefers to align the primary stress with the leftmost edge of the word (ω). This preference is also evident in her pronunciation of *Mexico City*, in which she placed the main stress on the leftmost syllable of each word: ω(Σ('mɛk.sɪ) Σ(ˌkoʊ)) ω(Σ('sɪ.tʰi)). Alternatively, she could have placed the main stress on [koʊ] but did not. Both examples show that the consultant consistently placed the main stress on the leftmost syllable of the word, regardless of the stress pattern of the native speaker.

Place names	Mandarin (Pinyin)	IPA (Target)	IPA (Consultant)
Mississippi		[ˌmɪ.sə.'sɪ.pi]	['mɪ.sɪ.ˌsɪ.pi]
密西西比	mi4 xi1 xi1 bi3	[mi^{51} ɕi^{55} ɕi^{55} pi^{21}]	[mi^{51} ɕi^{51} ɕi^{55} pi^{21}]
Mexico City		['mɛk.sɪ.ˌkoʊ 'sɪ.ɾi]	['mɛk.sɪ.ˌkoʊ 'sɪ.tʰi]
墨西哥 城	mo4 xi1 ge1 cheng2	[mo^{51} ɕi^{55} kɤ^{55} tʂʰəŋ^{35}]	[mo^{51} ɕi^{51} kɤ^{55} tʂʰəŋ^{35}]

Table 13. The stress and tone patterns of the consultant's production of place names in English and Mandarin.

4.3. Comparison of Mandarin, Cantonese, and English intonation in declarative questions

4.3.1 Raised pitch range

A raised pitch range is a specific intonation pattern that Mandarin uses to signal a question. This grammatical use of intonation is not a feature of Cantonese or English. Surprisingly, when the question preceded the statement in the reading (*i.e.*, Q-A order), the declarative questions in all three languages showed a higher pitch range than the statements. However, when the statement preceded the question (*i.e.*, S-Q order), only Mandarin and Cantonese questions showed this pitch effect. A third reading order in which the question preceded the answer but followed the statement (*i.e.*, S-Q-A order) confirmed evidence of a raised pitch range pattern in the consultant's Mandarin and Cantonese declarative questions but not in her English questions. The reason why the question had a higher pitch range than the statement in English in the Q-A order was due to the universal tendency for the F0 to gradually decrease toward the end of the discourse (Gussenhoven 2004). Thus, the initial sentence in a discourse is usually higher in pitch than the second sentence.

Figure 1 compares the consultant's production of English intonation patterns with Mandarin intonation patterns in declarative questions. Each pitch track contains a statement followed by a declarative echo question that repeats the entire statement. The Mandarin statement-and-question pairs each ends with a different Mandarin tone. The English sentences are translations of the Mandarin sentences.

In the English examples, the statement intonation shows an emphasis on the main verb *wants*, followed by a gradual fall. The question intonation begins at the F0 level of the tail end of the statement and gradually moves upward until near the end of the utterance where it rises sharply to signal a yes/no question. The raised pitch range pattern of Mandarin is not evident in the English questions.

In the Mandarin examples, the statement intonation shows an emphasis on the subject pronoun *ta1* 'he' instead of the first verb *xiang3* 'want', probably because *ta1* 'he' has

a high-level tone [55] and *xiang3* 'want' has a low-fall tone [21]. The question intonation shows a higher pitch range than the statement intonation. At the tail end of the question, F0 rises except when the final tone is mT4, in which case, F0 falls. This is a tonal effect due to the fact that the offset of mT1 [55], mT2 [35], and mT3 [214] all are relatively higher in pitch than the immediately preceding pitch value(s).

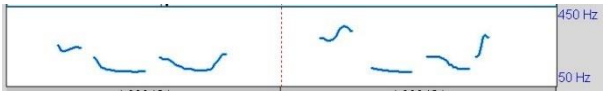
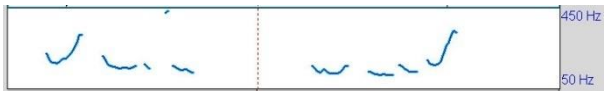
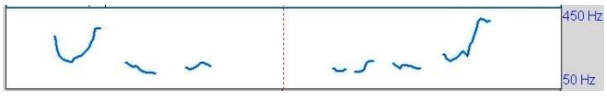
English	**Mandarin**
 (eS1) He wants to see his Mom. (eQ1) He wants to see his Mom?	 (mS1) 他想見他媽媽. (mQ1) 他想見他媽媽? *ta1 xiang3 jian4 ta1 ma1.ma* he want see he Mom 'He wants to see his Mom'
 (eS2) He wants to buy some hemp. (eQ2) He wants to buy some hemp?	 (mS2) 他想買大麻. (mQ2) 他想買大麻? *ta1 xiang3 mai3 da4.ma2* he want buy hemp 'He wants to buy some hemp'
 (eS3) He wants to ride a horse. (eQ3) He wants to ride a horse?	 (mS3) 他想騎馬. (mQ3) 他想騎馬? *ta1 xiang3 qi2 ma3* he want ride horse 'He wants to ride a horse'
 (eS4) He wants to be scolded. (eQ4) He wants to be scolded?	 (mS4) 他想被罵. (mQ4) 他想被罵? *ta1 xiang3 bei4 ma4* he want *(passive particle)* scold 'He wants to be scolded'

Figure 1. Examples of English and Mandarin statement and question intonation. In each pitch track, the F0 contour (shown in blue) to the left of the vertical dotted red line belongs to the statement, and the pitch contour to the right belongs to the question.

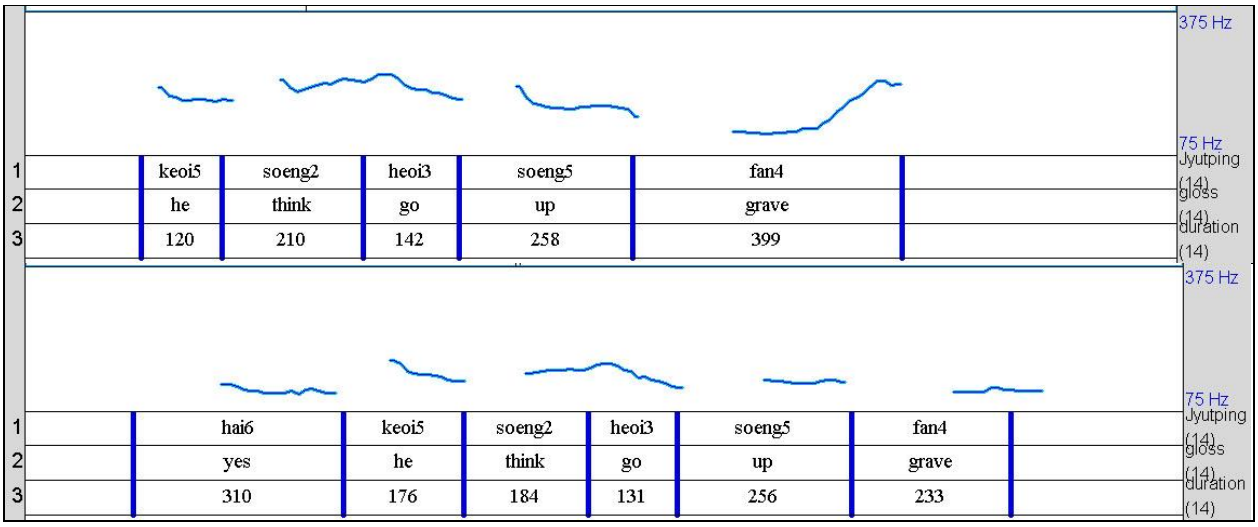


Figure 2. Pitch contours of a question and a statement in Cantonese:
'He wants to visit the grave? Yes, he wants to visit the grave.'

The raised pitch range effect of Mandarin is also evident in the consultant's production of Cantonese declarative questions, as shown in the example in Figure 2. This elevated intonation pattern is not native in Cantonese. Figure 3 shows a statement-question-answer sequence produced in Mandarin. The declarative question in the middle has a higher pitch throughout the utterance than the statement on the left and the answer-statement on the right. This is evidence of a raised pitch range effect on Mandarin questions. There also appears to be a lowered pitch effect on the last answer-statement, which is the declination effect described earlier.

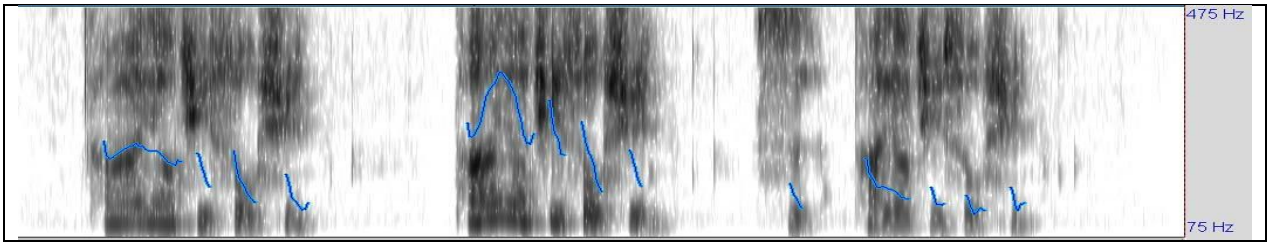


Figure 3. The pitch contours of the consultant's production of the sequence of a statement, question, and answer in Mandarin: *Ta1 yao4 qu4 shou4 xun4.* 'He must go for training.' *Ta1 yao4 qu4 shou4 xun4?* 'He must go for training?' *Shi4, ta1 yao4 qu4 shou4 xun4.* 'Yes, he must go for training.'

In summary, there is evidence of intonation transfer from Mandarin to Cantonese in the consultant's production of declarative questions. This intonation transfer is not evident in the English questions.

4.3.2 High boundary tone

A high boundary tone appears in a final rise of an utterance, *i.e.*, intonational phrase (IP). The consultant's Mandarin intonation shows a lack of a high boundary tone at the end of declarative questions. The IP-final tone seems to depend on the utterance-final lexical tone.

If it is a falling tone [51], the boundary tone is low (L%) and the tail end of the utterance falls; otherwise, the boundary tone is high (H%) and the tail end rises. The influence of lexical tone on the IP-final tone is shown in Table 14.

	Cantonese		Mandarin		English	
Token	Utterance-final Tone	IP-final Pitch Direction	Utterance-final Tone	IP-final Pitch Direction	Utterance-final Tone	IP-final Pitch Direction
Q11	cT1	↑	mT1	↑	n/a	↑
Q12	cT2	↑	mT3	↑	n/a	↑
Q13	cT3	↑	**mT4**	↓	n/a	↑
Q14	cT4	↑	mT2	↑	n/a	↑
Q15	cT5	↑	**mT4**	↓	n/a	↑
Q16	cT6	↑	**mT4**	↓	n/a	↑

Table 14. Pitch movement ('↑' rise or '↓' fall) at the end of questions ending in 'fan' for Cantonese, and their corresponding sentences in Mandarin and English. English has no lexical tone (n/a).

The English declarative questions produced by the consultant show an upward pitch movement at the tail end, as indicated in Table 14. This intonation pattern is consistent with that of a native speaker of English.

Finally, in the case of Cantonese declarative questions, the consultant produced them with a rising F0 at the end regardless of the lexical tone at utterance-final position. This rising intonation pattern is consistent with the question pattern produced by native speakers of Cantonese. Figures (4, 5) show examples of the consultant's production of a question ending in cT5 and cT6, respectively. In both cases, the boundary tones were H%.

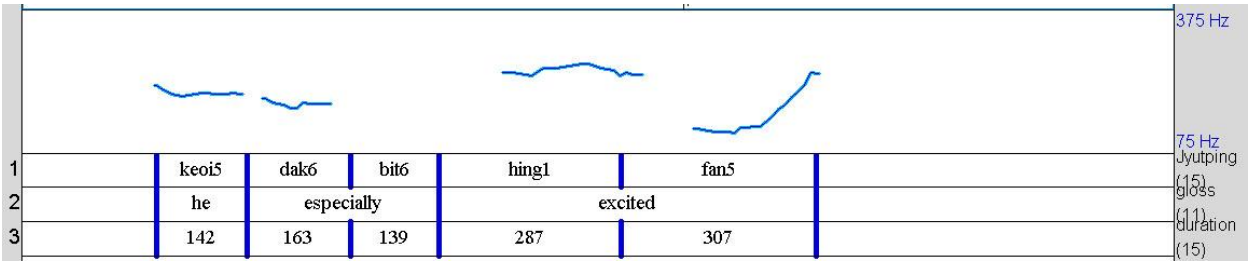


Figure 4. The consultant's production of a question ending in cT5 in Cantonese: *Keoi5 dak6.bit6 hing1.fan5?* 'He is especially excited?' (Only partial annotation is shown.)

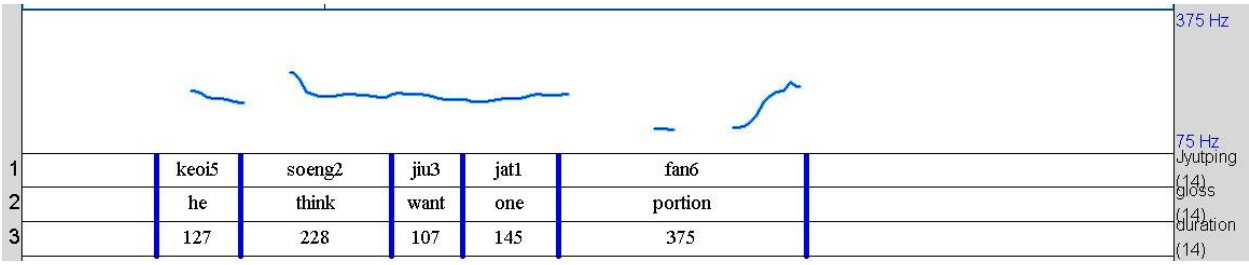


Figure 5. The consultant's production of a question ending in cT6 in Cantonese: K*eoi5 soeng2 jiu3 jat1 fan6?* 'He wants to have a portion?' (Only partial annotation is shown.)

5. Conclusion

This study examined the speech of a native speaker of Mandarin to determine if there was an L2 transfer of intonation, as well as lexical tones and word stress, from Mandarin to Cantonese and English. As predicted, the consultant demonstrated difficulty in producing the low register tones in Cantonese (*i.e.*, cT5 and cT6) due to their markedness as compared to Mandarin tones. It was also predicted that there would be influence from her Mandarin stress pattern on her English. Her preference for a trochaic stressed binary foot was evident in her production of Mandarin disyllabic words and place names. Her production of English disyllabic words also consistently showed a pattern of trochees and a left-headed word, even when an alternate, acceptable pattern was possible.

From the elicited speech samples, I was unable to determine if the stress pattern of Mandarin had an effect on the consultant's English sentence intonation. However, the intonation on her English questions was near native-like and did not show any signs of intonation transfer from Mandarin. Cantonese, on the other hand, was affected by Mandarin. Although the consultant showed acquisition of the native Cantonese question pattern, she also adopted the raised pitch range effect from her Mandarin questions. The result of this study suggests that the consultant encountered more difficulty during her acquisition of the intonation of Cantonese than that of English. The implication is that lexical tones interfered with sentence intonation more so than word stress for this speaker. This is in accordance with Yuan's (2011) claim that both lexical tones and intonation use pitch as a primary cue.

This study examined the speech of a single speaker from Singapore. A future direction would be to repeat the study with other speakers, including males, teenagers, and speakers of other dialects of Mandarin (*e.g.*, Taiwan's Guoyu and China's Putonghua).

This case study shows that lexical tones are difficult to acquire, even for an L2 learner who has acquired the lexical tones of her L1. It also shows that word stress is difficult to acquire, especially in an L2 language, such as English, which has complex stress patterns.

One related finding from this study is the potential link between cT1 and cT6 in Cantonese and syllable stress in English. In English and Cantonese loan words, such as *strawberry* and *bus*, a stressed syllable (S) is associated with a [55] tone (cT1) and a non-stressed syllable (NS) is associated with a [22] tone (cT6), as shown in Table 15.

strawberry	bus
[s ˈtɹɑ ˌbɛ ɹi] ↕ ↕ ↕ [siː^{22} tɔː^{55} pɛː^{55} lei^{25}] NS S S	[ˈbʌ s] ↕ ↕ [paː^{55} siː^{22}] S NS

Table 15. Examples showing a link between a stressed syllable (S) and a [55] tone, and a link between a non-stressed syllable (NS) and a [22] tone.

This characteristic is also seen in S-language (Jernudd & Yan 1995), a secret language used by teenage students in Hong Kong, in which the tone of the actual syllable is changed to cT6

(to make it the least prominent) and the reduplicated syllable which follows is assigned cT1 (to make it the most prominent). This phenomenon, in which a high tone is associated with stress and a low tone is associated with no stress, is perhaps unique to Hong Kong Cantonese because of its linguistic interaction with the English language, and perhaps specific to loan words.

According to Flynn (2003), although there is no contrast of stressed and unstressed syllables in Cantonese, there are varying degrees of prominence between syllables. If one important element in creating prominence in Cantonese is lexical tones (complemented by duration and, optionally, intensity), could there be common perception cues between how prominence is achieved in Cantonese and how stress is realized in English? Could these common cues help an L2 learner to perceive these prosodic elements in their non-native language? Future research in this area might provide insight into L2 English stress patterns of Hong Kong Cantonese speakers.

References

Bauer, Robert S. & Paul K. Benedict. 1997. *Modern Cantonese Phonology*. Trends in Linguistics Studies and Monographs 102. Berlin: Mouton de Gruyter.

Beckman, Mary E., Julia Hirschberg, & Stefanie Shattuck-Hufnagel. 2005. The original ToBI system and the evolution of the ToBI framework. In S.-A. Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*. 9-54. New York: Oxford University Press.

Boersma, Paul & David Weenink. 2010. *Praat: Doing phonetics by computer*. (Computer application, version 5.1.43). Downloaded from http://www.praat.org, 2010-08-04.

Chao, Yuen-Ren. 1930. A system of tone letters. *Le Maitre Phonetique* 30. 24-27.

Chen, Xi, Richard C. Anderson, Wenling Li, Meiling Hao, Xinchun Wu, & Hua Shu. 2004. Phonological awareness of bilingual and monolingual Chinese children. *Journal of Educational Psychology* 96(1). 142-151.

Ciocca, Valter, Alexander L. Francis, Rani Aisha, & Lena Wong. 2002. The perception of Cantonese lexical tones by early-deafened cochlear implantees. *Journal of Acoustical Society of America* 111. 2250-2256.

Duanmu, San. 2007. *The phonology of standard Chinese, 2nd edn*. Oxford: Oxford University Press.

Eckman, Fred R. 1997. Markedness and the contrastive analysis hypothesis. *Language Learning* 27. 315-330.

Flynn, Choi Yeung Chang. 2003. *Intonation in Cantonese*. LINCOM Studies in Asian Linguistics 49.

Gu, Wentao, Keikichi Hirose, & Hiroya Fujisaki. 2005. Analysis of the effects of word emphasis and echo questions on F0 contours of Cantonese utterances. *Interspeech 2005*. 1825-1828.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Jernudd, Björn H. & Eleanor Y. Y. Yan. 1995. A secret language in Hong Kong. *Cahiers de linguistique – Asie orientale* 24(1). 155-166.

Ladd, D. Robert. 2008. *Intonational phonology, 2nd edn.* Cambridge: Cambridge University Press.

Marinova-Todd, Stefka H., Jing Zhao, & May Bernhardt. 2010. Phonological awareness skills in the two languages of Mandarin-English bilingual children. *Clinical Linguistics & Phonetics* 24(4-5). 387-400.

Mok, Peggy. 2009. On the syllable-timing of Cantonese and Beijing Mandarin. *Chinese Journal of Phonetics* 2. 148-154.

Peng, Shu-Hui, Marjorie K. M. Chan, Chiu-Yu Tseng, Tsan Huang, Ok Joo Lee, & Mary E. Beckman. 2005. Towards a Pan-Mandarin system for prosodic transcription. In S.-A. Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing.* 230-270. New York: Oxford University Press.

Pulleyblank, Douglas. 1986. *Tone in lexical phonology*. Dordrecht: D. Reidel.

Wells, John C. 2006. *English intonation: An introduction*. Cambridge: Cambridge University Press.

Williams, Carl E. & Kenneth N. Stevens. 1972. Emotions and speech: Some acoustical correlates. *Journal of the Acoustical Society of America* 52(4). 1238-1250.

Wong, Wai Yi P., Marjorie K. M. Chan, & Mary E. Beckman. 2005. An autosegmental-metrical analysis and prosodic annotation conventions for Cantonese. In S.-A. Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*. 271-300. New York: Oxford University Press.

Yip, Moira. 1980. *The tonal phonology of Chinese*. PhD dissertation, Department of Linguistics, MIT.

Yuan, Jiahong. 2011. Perception of intonation in Mandarin Chinese. *Journal of the Acoustical Society of America* 130(6). 4063-4069.

Yuan, Jiahong, Chilin Shih, & Greg P. Kochanski. 2002. Comparison of declarative and interrogative intonation in Chinese. *Proceedings of Speech Prosody 2002*. 711-714.

Appendix A: MC_ToBI

Tier	Description
tones	lexical tones and intonational phrase boundary tones
stress	stress levels associated with a full, reduced or neutral tone in Mandarin
breaks	end of the syllable, word, intermediate phrase or intonational phrase
pitch range	change or reset in pitch range
foot	(a) fusion forms and their combined tones, *e.g.*, jik5+hai22 → je55-22 'that is' (b) emphasized syllables and phrases (*e.g.*, longer duration, higher intensity, expanded pitch range relative to any pitch range change noted in the *pitch range* tier), tagged with **<*>**, or **<*** and ***>**.
syllables	(a) romanization of the Chinese character in the *words* tier: Jyutping for Cantonese and Pinyin for Mandarin (b) intervals of silence, labeled with <SIL>
words	Chinese character
gloss	English word-by-word translation
misc	(a) sentence type: <Q> for question and <S> for statement (b) the correct or normal lexical tone, if it is different from the transcribed tone

Table A1. MC_ToBI: Tiers, derived from M_ToBI (Peng *et al.* 2005) and C_ToBI (Wong *et al.* 2005)

Lexical Tones			Cantonese			Mandarin	
55	high-level		cT1	55		mT1	55
33	mid-level		cT3	33			
22	low-level		cT6	22			
5	high-level	(σ-final stop)	cT1	5	(for checked syllable)		
3	mid-level	(σ-final stop)	cT3	3	(for checked syllable)		
2	low-level	(σ-final stop)	cT6	2	(for checked syllable)		
23	low-rise		cT5	23			
35	high-rise		cT2	25 or 35		mT2	35
51	high-fall					mT4	51
53	high-mid-fall		cT1	53	(allophone of 55)		
32	mid-fall						
21	low-fall		cT4	21		mT3	21
214	low-fall-rise					mT3	214

Table A2. MC_ToBI: Tone numbers, and their equivalent in Cantonese and Mandarin.
The mid-fall [32] tone is not a specified Mandarin or Cantonese tone but appears in the elicited data.

Boundary Tone	Description
L%	fall from the IP-final lexical tone
H%	rise from the IP-final lexical tone
LH%	fall then rise from the IP final lexical tone
HL%	rise then fall from the IP final lexical tone
%	no extra tone added to the IP-final lexical tone

Table A3. MC_ToBI: Intonational phrase (IP) boundary tones, from C_ToBI (Wong *et al.* 2005:296)
with 'LH%' added

Stress Level	Description
S3	syllable with fully-realized lexical tone
S2	syllable with substantial tone reduction
S1	syllable that has lost its lexical tonal specification
S0	syllable with lexical neutral tone

Table A4. MC_ToBI: Stress levels in the *stress* tier, from M_ToBI
(These labels and descriptions were extracted from Peng *et al.* 2005:265.)

Break Indice	Description
0	reduced syllable boundary, or syllable boundary of a fused form
1-	syllable boundary within a polysyllabic word
1	word boundary
2	uncertainty between two boundaries or mismatch
3	intermediate phrase boundary (with a perceived pause or prolongation)
4	intonational phrase boundary (with reset of pitch) which is associated with a boundary tone on the *tones* tier

Table A5. MC_ToBI: Break indices, derived from MAE_ToBI (Beckman *et al*. 2005:23), C_ToBI (Wong *et al.* 2005:297), and M_ToBI (Peng *et al.* 2005:266)

Tag	Description
%reset	beginning of a new pitch downtrend or pitch reset
%q-raise	beginning of a raised pitch range (*e.g.*, in echo questions)
%e-prom	beginning of local expansion of pitch range due to emphatic prominence
%compressed	beginning of reduction of pitch range of syllables following the expansion of pitch range under %e-prom

Table A6. MC_ToBI: Tags for pitch range effects in the *pitch range* tier, from M_ToBI
(These tags and descriptions were extracted from Peng *et al.* 2005:260.)

Contact Information:

Una Y. Chow

uchow@ucalgary.ca

School of Languages, Linguistics, Literatures & Cultures
The University of Calgary
2500 University Drive N.W.
Calgary, AB, T2N 1N4
Canada